

University of Groningen

A dynamic spatial econometric diffusion model with common factors

Ciccarelli, Carlo; Elhorst, J. Paul

Published in:
Regional Science and Urban Economics

DOI:
[10.1016/j.regsciurbeco.2017.07.003](https://doi.org/10.1016/j.regsciurbeco.2017.07.003)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Ciccarelli, C., & Elhorst, J. P. (2018). A dynamic spatial econometric diffusion model with common factors: The rise and spread of cigarette consumption in Italy. *Regional Science and Urban Economics*, 72, 131-142. <https://doi.org/10.1016/j.regsciurbeco.2017.07.003>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A dynamic spatial econometric diffusion model with common factors: the rise and spread of cigarette consumption in Italy *

Carlo Ciccarelli[†] J.Paul Elhorst[‡]

July 3, 2017

Abstract

This paper adopts a dynamic spatial panel data model with common factors to explain the non-stationary diffusion process of cigarette consumption across 69 Italian provinces over the period 1877-1913. The CD-test of Pesaran (2015a), the exponent α -test of Bailey et al. (2015), the cross-sectionally augmented panel unit root test of Pesaran et al. (2013), and the spatial stability test of Yu et al. (2012) are used to show that both global common factors and local spatial dependence are important drivers of the propagation of cigarette demand over this period and to determine the point at which the hypothesis of stationarity no longer needs to be rejected in favor of a unit root. The direct and indirect effects derived from the coefficient estimates of the model show that cigarettes were a normal good with an income elasticity of 0.4 and a price elasticity -0.4 in the long term. This price elasticity of -0.4 consists of a direct effect of -0.54 in the own region and a spillover effect to other regions of 0.15. This positive spillover effect is in line with previous spatial econometric studies which investigated cigarette demand in the U.S. states over a more recent period.

Keywords: diffusion, stationarity, spatial dependence, common factors, cigarette demand

JEL Classification: C21, C23, N33, N93, R22

1 Introduction

This paper sets out a dynamic spatial panel data model with common factors to explain cigarette diffusion in Italy over the period 1877-1913. The observations

*We thank Roberto Basile, Roberto Benedetti, Luisa Corrado, Guido De Blasio, Valter di Giacinto, Tarek Harchaoui, Alain Pirotte, two anonymous reviewers, and the editor Zhenlin Yang for valuable comments and suggestions. The usual disclaimer applies.

[†]Department of Economics and Finance, University of Rome Tor Vergata, Via Columbia 2, 00133 Rome, Italy. *Email*: carlo.cicarelli@uniroma2.it

[‡]Faculty of Economics and Business, University of Groningen. Nettelbosje 2, 9747 AE Groningen, The Netherlands. *Email*: j.p.elhorst@rug.nl

that cigarettes represent a new and trendy product at the turn of the 19th century in Italy and that the data over this period are non-stationary make the modeling of its diffusion process an interesting and challenging topic of research. The proposed model will be presented in a such a general form that it also can be used to analyze other and more recent diffusion processes.

Although tobacco was a well-established industrial sector in our study period (1877-1913), cigarettes represented an entirely new commodity for Italian consumers. After an introductory stage, in which the size of the market for cigarettes was still very small, policy makers implemented several measures to increase its diffusion among the Italian population, with the aim to increase public revenues. The tobacco sector in Italy is very well documented. The empirical evidence presented in this paper is based on primary historical sources, specifically the annual budget reports of the tobacco companies. The data, to be discussed in more detail in Section 3, are characterized by an initial period of extremely low levels of per capita consumption, followed by a rapid increase and acceleration concentrated in a short period of time. The availability of these data at the provincial level (NUTS 3 level) allows us to consider the propagation of cigarette consumption not only over time but also across space.

The proposed model is both dynamic and spatial. It is dynamic in that we account for serial dependence across the data in each province over time, and spatial in that we account for both global common factors and local spatial dependence across the observations at each point in time, also known as strong and weak cross-sectional dependence in the literature (Chudik et al., 2011). Standard econometric test procedures, the CD-test developed by Pesaran (2004, 2015a), point to cross-sectional dependence in cigarettes consumption, while more recent testing procedures, the exponent α -test of Bailey et al. (2015), offers the opportunity to test for weak against strong cross sectional dependence or to test for strong cross sectional dependence first, and then after common factors have taken up for weak cross-sectional dependence. For cases where both are likely to be present, Bailey et al. (2016) developed a two-stage estimation procedure first to address strong and then to address weak cross-sectional dependence, while Halleck-Vega and Elhorst (2016) developed an approach that simultaneously accounts for both forms of cross-sectional dependence, as well as serial dynamics, within one framework. However, both studies did not include any independent variables. We demonstrate that this simultaneous framework extended to include such variables covers several simpler spatial econometric models that have been considered in the literature before and that it provides an adequate tool to describe diffusion processes in general, and the diffusion process of cigarette consumption in Italy in particular. Finally, this study is among the first to consider the sensitivity of direct and indirect effects estimates of the explanatory variables, which can be derived from the coefficient estimates of the model, for the inclusion of common factors. Cigarettes are found to be a normal good with an income elasticity of 0.4 and a price elasticity -0.4 in the long term. This price elasticity can be further decomposed into a direct effect of -0.54 in the own region and a spillover effect to other regions of 0.15.

The paper is organized as follows. Section 2 provides background on the

diffusion of cigarettes in Italy, introduces the proposed empirical model, discusses related literature, and presents a series of tests to examine whether serial dependence, spatial dependence and common factors have been modeled effectively such that the model is stable. Section 3 describes the data and its main sources, and illustrates the main spatiotemporal patterns of per capita cigarette consumption in Italy’s provinces. Section 4 reports and discusses the main empirical findings, while Section 5 concludes.

2 Modeling strategy and review of the literature

After providing the necessary background on the early diffusion of cigarettes consumption in Italy, this section illustrates the proposed empirical model and discusses related literature.

2.1 Background on the diffusion of cigarettes in Italy

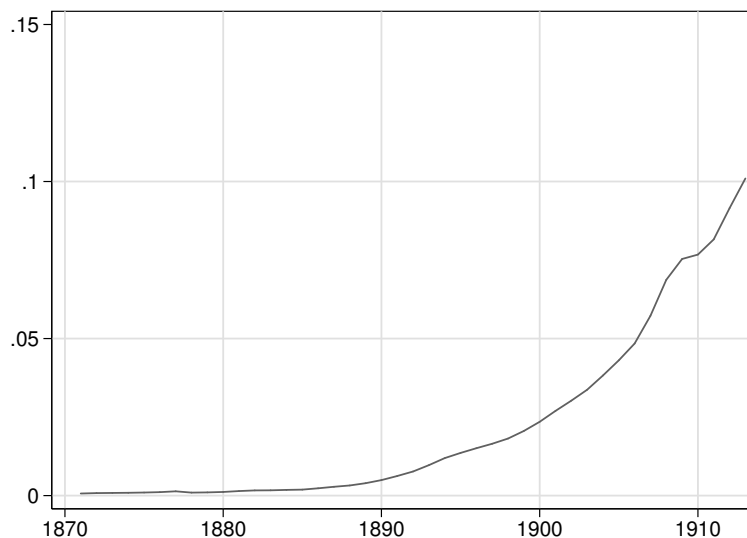
Historical studies document that most Italian tobacco factories were active in the 18th century; the tobacco factory in Turin started operating in 1740, the one in Florence in 1769 and the one in Milan in 1771 (Cappellari della Colomba, 1866). Soon after its political unification in 1861, the Italian State established a public monopoly with roughly one factory assigned to each region, and with the same set of rules (production process, selling price, wages) extended to the whole Italian territory, so as to raise more revenues from the sales of tobacco for the public budget.¹ Already in 1865, this resulted in 15 public tobacco factories with more than 14,000 workers (mostly women). Health related issues were almost completely ignored at that time.² The replacement of hand-rolled by machine-rolled cigarettes, both inside and outside Italy, contributed to a reduction in the cost of production and subsequently to a fall of the price of cigarettes. The generalized increase of literacy rates, the beginning of mass-press, the diffusion of railways, the increasing process of urbanization and all that, rapidly increased the circulation of commodities and ideas, and contributed to the emergence of a new social behavior and life style of the Italians. In addition, and most importantly, income increased rapidly during the *belle époque*, roughly the period from 1895 to 1913. Indeed, historical studies show a clear increase in per capita consumption of cotton and wool, but also of beer, coffee, and sugar in those years (Fenoaltea 2011, p. 120). Various circumstances contributed thus to what we might call the “territorial take-off” of cigarette consumption in Italy

¹As reported in Ragioneria generale dello Stato (1914, pp. 424, 426, 448, 450), in 1871, at the beginning of our sample period, public revenues from total tobacco sales amounted to 72.93 million lire, corresponding to some eight per cent of total ordinary public revenues (945.47 million lire). In 1912-13, at the end of our sample period, public revenues from total tobacco sales amounted to 333.06 million lire, corresponding to some 13 per cent of total ordinary public revenues (2,491.95 million lire).

²We note however that Scalzi (1868), in his study on the consumption of tobacco in Rome, analyzed the negative consequences of smoking tobacco on health. Interestingly, when computing the potential population of smokers in Rome in 1866, he excluded youngsters aged 0-14 (23,814), women (99,892), and clergymen (5,209) from the total population (210,701).

at the turn of 19th century. As Figure 1 illustrates, per capita consumption of cigarettes in physical terms followed an exponential growth path over the period 1871-1913. After the 1870s and the 1880s in which the level of consumption was still relatively low, consumption started to rise, to reach a peak of about 0.1 kg or 100 cigarettes in 1913, the last year of our sample period. The present study provides an attempt to model this diffusion process empirically.

Figure 1. Per capita consumption of cigarettes (kg), 1871-1913^a



^a Data for Sicilian provinces not available during 1871-1876.

The temporal pattern illustrated in Figure 1 for Italy is in line with Rogozinsky (1990, p. 166) who noticed that in many countries “Cigarettes [...] become popular among smokers only starting in the 1880s with the introduction of machine-made brands”. According to Diana (1998, pp. 44-45), cigarettes were introduced in Italy by the veterans of the Crimean War (1853-1856),³ while according to Cappellari Della Colomba (1865, p. 345), the tobacco factory in Florence was the only one in Italy producing cigarettes in 1865 (the production of cigarettes in 1865 amounted to 4,345 kilograms). In addition, Pasetti (1906) reports that “total production of tobacco declined during 1888-89 of some 2% when compared to the previous year [...] however it was sustained by the rising consumption of cigarettes produced in Italy” (p. 69), and, when referring to 1890-91 “In this year in almost any tobacco factory the machinery was renewed”

³To get support from foreign countries in the process of national unification (culminated with the independence wars and the subsequent political unification of Italy in 1861), the Kingdom of Sardinia decided to send about 15,000 soldiers to side with French and British forces.

(p. 73), and, when referring to 1891-92, “there was an extraordinary increase in the production of cigarettes” (p. 76). It is estimated that a well-trained worker was able to roll about 1,000-1,200 cigarettes during a standard (eight hours) working day. Since the early 1890s, machine rolled cigarettes were effectively produced and sold in Italy. Six new Bonsack machines were installed in the tobacco factory of Rome (Ministero delle finanze, 1893, p. XXIII).⁴ The Bonsack machines were rather rudimentary when compared to the subsequent machines in the twentieth century. Still, they were able to roll about 200 cigarettes per minute, corresponding to some 96 thousands per working day. The cost reduction implied by technological progress was substantial, and cigarettes started to be an affordable and ordinary consumption good in Italy. In addition, as noticed in Cross and Proctor (2014, pp. 69-70), matches also helped popularize cigarettes by “increasing the ease and convenience [and safety] of making fire, allowing a quick and calibrated ignition.”

Historical studies on tobacco are relatively abundant. Madson (1916) considers the production and consumption of tobacco in various countries where the tobacco business was also run by the State according to a regime of public monopoly (among which France, Italy, Austria, Japan, Spain, Sweden). Hutson (1937) provides a similar detailed analysis of tobacco consumption in European countries during the period 1913 to 1937. Historical studies on tobacco in Italy include, among others, Cappellari della Colomba (1866) who provides an historical account on the manufacturing of tobacco in Italy since the 18th century, and Pasetti (1906) with a concise yet useful account of the annual evolution of the tobacco sector from 1884 to 1905. More recent contributions use econometric techniques to analyze data on tobacco consumption. Manera (1963) and Ciccarelli, Pierani and Tiezzi (2014) use national data on the consumption of cigars, cut tobacco, snuff tobacco, and cigarettes. Both studies follow a time series approach. Ciccarelli, Pierani and Tiezzi (2014) show in particular that the aggregate consumption of cigarettes in Italy increased for decades, peaked in the 1980s and then started to decrease when people became aware of the negative consequences of smoking for their health. Ciccarelli and De Fraja (2014) use historical provincial data on tobacco consumption based on the same historical sources as in the present study, but they focus on addiction and its main determinants rather than diffusion. Most importantly, in each of the aforementioned studies on tobacco consumption in Italy the impact of cross-sectional dependence is ignored altogether.

2.2 The econometric specification

The dynamic spatial panel data model with common factors that is taken as point of departure in this paper reads as

⁴J. A. Bonsack, is widely credited to be the inventor in 1880 of the first cigarettes rolling machine (“Cigarette-Machine - US Patent 238,460. September 4, 1880”). Diana (1999, p. 28) mentions the visit of J.A. Bonsack to the tobacco factory of Rome in 1882 to promote his new cigarettes rolling machine.

$$\begin{aligned}
C_t = & \tau C_{t-1} + \delta W C_t + \eta W C_{t-1} + X_t \beta + W X_t \theta \\
& + \Gamma_1 \bar{C}_t + \Gamma_2 \bar{C}_{t-1} + \sum_{k=1}^K \Pi_k \bar{X}_{kt} + \mu + \xi_t \iota_N + \epsilon_t
\end{aligned} \tag{1}$$

where C_t denotes an $N \times 1$ column vector that consists of one observation of the dependent variable for every unit ($i = 1, \dots, N$) in the sample at time t ($t = 1, \dots, T$), which for this study is per capita consumption of cigarettes in the 69 Italian provinces over the period $t = 1877, 1878, \dots, 1913$. X_t represents an $N \times K$ matrix of independent variables, which for this study are the real price of cigarettes, real income, and the literacy rate. All variables are measured in logs. C_{t-1} and $W C_t$ represent, respectively, the temporal and spatial lag, and $W C_{t-1}$ the spatiotemporal lag of C_t , while τ , δ , and η are the response parameters of these variables, better known as, respectively, the serial, spatial, and spatiotemporal autoregressive coefficients. The $N \times N$ matrix W is a non-negative matrix of known constants describing the spatial arrangement of the spatial units in the sample. The specification of this matrix will be further discussed in the empirical application. Since the spatial econometric model in Equation (1) contains both X and $W X$ variables, it is also known as a dynamic spatial Durbin model (Elhorst, 2014, p. 106). The variables described so far are meant to cover potential spatial dependence (weak cross-sectional dependence) among the observations. The common factor terms meant to cover potential strong cross-sectional dependence are defined as the cross-sectional averages of the dependent variable at times t and $t-1$ or of the independent variable at time t , i.e. $\bar{C}_t = \frac{1}{N} \sum_{i=1}^N C_{it}$, $\bar{C}_{t-1} = \frac{1}{N} \sum_{i=1}^N C_{it-1}$, and $\bar{X}_{kt} = \frac{1}{N} \sum_{i=1}^N X_{ikt}$, where k denotes the k^{th} independent variable of \bar{X}_{kt} . Alternatively, these common factors may be defined as the population weighted and thus national averages across the provinces in the sample.⁵ It is important to stress that, while the variables covering weak spatial dependence have common response parameters, the common factors enter the equation with unit-specific coefficients stored in the $N \times 1$ column vectors Γ_1 , Γ_2 , and Π_k for $k = 1, \dots, K$. The original idea to control for common factors in a non-dynamic and non-spatial model goes back to Pesaran (2006). Pesaran et al. (2013) extends this idea to a dynamic but non-spatial model, Bailey et al. (2016) to a dynamic and spatial model, though without any independent variables and by addressing strong and weak cross-sectional dependence in two separate stages. Finally, Halleck-Vega and Elhorst (2016) demonstrate that if both types of cross-sectional dependence are accounted for simultaneously not only the cross-sectional average of the dependent variable at time t but also at time $t-1$ needs to be controlled for, just as in Pesaran et al. (2013). Since the numbers of parameters increases rapidly with the number of common factors, every common factor requires N

⁵The terminology common factors should not be confused with the common factor test originally proposed by Burridge (1981), and further analysed by Mur and Angulo (2006), to examine whether the spatial Durbin model can be simplified to the spatial error model.

additional parameters to be estimated, this paper tests different sets of common factors against each other. The vectors $\mu = (\mu_1, \mu_2, \dots, \mu_N)'$, and $\xi_t \iota_N$, where ι_N is an $N \times 1$ column vector of ones, represent, respectively, spatial fixed effects and time-period fixed effects. These fixed effects are optional since they do not always go together with common factors, as will be discussed later. Finally, the $N \times 1$ vector $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{Nt})'$ consists of i.i.d. disturbance terms, which have zero mean and finite variance σ^2 .

The parameters of Equation 1 can be estimated by the bias-corrected maximum likelihood estimator developed by Yu et al. (2008) when spatial fixed effects are controlled for, and by Lee and Yu (2010a) if, in addition, time period fixed effects are included. The common factors may be treated as exogenous explanatory variables based on the assumption that the contribution of each single province to the cross-sectional averages at a particular point in time goes to zero if N goes to infinity (Pesaran, 2006, assumption 5 and remark 3). Since we use data of 69 provinces this assumption is more likely to be met than in Halleck-Vega and Elhorst (2016), whose analysis is based on 12 provinces only.

The dynamic spatial panel data model without the inclusion of common factors has gained a lot of attention in the spatial econometrics literature, in terms of estimation (ML, GMM/IV and Bayesian MCMC), in terms of interpretation, and in terms of specific applications. Two recent overviews of this literature are provided by Lee and Yu (2010b) and Elhorst (2014, Ch. 4). Direct interpretation of the coefficients in Equation (1) is difficult, because they do not represent the marginal effect of the independent variables. Elhorst (2014, Ch. 4) shows that the matrix of marginal effects of the expected value of the dependent variable with respect to the k th independent variable in unit 1 up to unit N (say x_{ik} for $i = 1, \dots, N$, respectively) in the long term is given by the $N \times N$ matrix

$$\begin{aligned} \left(\begin{array}{ccc} \frac{\partial E(C_t)}{\partial x_{1k}} & \dots & \frac{\partial E(C_t)}{\partial x_{Nk}} \end{array} \right) &= \left(\begin{array}{ccc} \frac{\partial E(c_{1t})}{\partial x_{1k}} & \dots & \frac{\partial E(c_{1t})}{\partial x_{Nk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial E(c_{Nt})}{\partial x_{1k}} & \dots & \frac{\partial E(c_{Nt})}{\partial x_{Nk}} \end{array} \right) \\ &= ((1 - \tau)I_N - (\delta + \eta)W)^{-1} \left(\begin{array}{cccc} \beta_k & \theta w_{12} & \dots & \theta w_{1N} \\ \theta_k w_{21} & \beta_k & \dots & \theta w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_k w_{N1} & \theta_k w_{N2} & \dots & \beta_k \end{array} \right) \end{aligned} \quad (2)$$

whose average diagonal element can be used as a summary indicator for the direct effect, and average row sum of off-diagonal elements as a summary indicator of the spillover effect. Note that the full $N \times N$ matrix is the product of two $N \times N$ matrices. The elements of the first of these two matrices, the inverse of the matrix $(1 - \tau)I_N - (\delta + \eta)W$, better known as the spatial multiplier matrix, are not specified since their analytical expressions are unknown. The two mentioned summary indicators reflect the long-term impact on the dependent variable that result from a change in the k th independent variable x_k respectively in the own province and in other provinces. Their short-term

counterparts can be obtained by setting $\tau = \eta = 0$, while the significance levels of these short and long-term direct and spillover effects are bootstrapped (see Elhorst, 2014, Section 2.7.2 for details).⁶

Applications of the dynamic spatial panel data model from the very beginning are of Elhorst et al. (2010) on economic convergence, Parent and LeSage (2010) on commuting, and Brady (2011) on housing prices. However, neither of these studies considered direct and indirect effects estimates. The first studies reporting these effects are of Debarsy et al. (2012) and Elhorst (2014). The first studies incorporating common factors are of Bailey et al. (2016), Halleck-Vega and Elhorst (2016), Ertur and Musolesi (2016) and Carrion-i-Silvestre and Surdeanu (2016), but these studies on their turn do not consider direct and indirect effects estimates, since they do not include independent variables (first two studies) or only include spatial lags in the error term specification (last two studies). This study is therefore among the first to consider the sensitivity of direct and indirect effects estimates of the independent variables for the inclusion of common factors.

To conclude this section, we note that while the literature on the early steps of tobacco consumption in Italy largely overlooked the geographical dimension of cigarettes consumption, spatial dependence attracted a lot of attention in a series of studies based on the well-known data set on cigarette demand of Baltagi and Li (2004); a panel data set of 46 U.S. states over the period 1963 to 1992 that is often used for illustration purposes. This data set was used for the first time by Baltagi and Levin (1986, 1992), but then respectively over the periods 1963-1980 and 1963-1988. All other studies mentioned in this paragraph utilized the full data set. Today most studies control for spatial and time period fixed effects (Elhorst, 2014; Kelejian and Piras, 2014; Halleck-Vega and Elhorst, 2015). Elhorst (2014) explicitly tests for these controls and finds that this model specification outperforms its counterparts without spatial and/or time fixed effects, as well as the random effects model. Many studies also include the dependent variable lagged in time to control for habit persistence, leading to the dynamic spatial panel data model (Baltagi and Levin, 1986, 1992; Elhorst, 2014; Debarsy et al., 2014). In that case one can distinguish both short-term and long-term direct and indirect effects using Equation 2. Most studies also share the view that spatially lagged independent variables (WX) should be included (Baltagi and Levin, 1986, 1992; Elhorst, 2014; Debarsy et al., 2014). Finally, almost all studies adopt a row-normalized binary contiguity matrix. One exception is Debarsy et al. (2014) who also consider a row-normalized matrix based on state border miles in common between states. However, up to now, not one single study based on this data set controlled for common factors.

⁶In contrast to the spatial econometrics literature, the (G)VAR literature is generally more focused on the impact of idiosyncratic shocks to the dependent variable in a given area on that of the area itself and on neighboring areas, where the impact of neighboring areas is sometimes labeled contagion. These effects can be simulated by replacing the second $N \times N$ matrix on the right-hand side of Equation 2 by a $N \times 1$ vector $S = (\dots, s_i, \dots)$, where s_i is generally set to one standard error of the error term representing the shock, and multiplying this shock by the spatial multiplier matrix.

2.3 Time-space recursive alternative

A related model that has gained a lot of attention in the spatial econometrics literature is the time-space recursive spatial econometric model. This spatial econometric model is similar to the dynamic spatial panel data model (without common factors), except that the spatially lagged dependent variable WC_t has been removed. According to Anselin et al. (2008), this model is especially useful to study spatial diffusion phenomena. LeSage and Pace (2009, ch. 7) refer to this model as a classic spatiotemporal (partial adjustment) model and use it to show that high temporal dependence and low spatial dependence might nonetheless imply a long-run equilibrium with high spatial dependence. Fogli and Veldkamp (2011) adopt this model to investigate whether the labor force participation decision can vary with past participation behavior in surrounding regions, based on decennial data of female participation rates over the period 1940-2000 at the U.S. county level. A similar study based on a panel of 108 regions across eight EU countries over the period 1986-2010 is carried out by Halleck-Vega and Elhorst (2017). Korniotos (2010) applies this model to explain annual consumption growth in U.S. states over the period 1966-1998 and interprets the coefficients of the temporal and space-time lags of the dependent variable (C_{t-1} and WC_{t-1}) as measures of what he calls the relative strength of internal and external habit persistence. This model and terminology is also used by Verhelst and Van den Poel (2014) in a similar study on daily transactions of six different stores of an European retailer from January 2002 until November 2004. A similar type of model is used in the marketing literature by Bollinger and Gillingham (2012) to explain the diffusion of solar panels. However, instead of considering C_{t-1} in the regression equation itself, they allow for a serial lag in the error term specification. More applications in the marketing literature are summarized in Elhorst (2017).

Despite the popularity of the time-space recursive spatial econometric model, a basic question is whether the removal of the spatially lagged dependent variable WC_t is supported by the data. Indeed, some researchers are troubled with the idea that the spatial autoregressive interaction between C and WC is instantaneous (see Upton and Fingleton, 1985, p. 369 for one of the first discussions on this issue). Instead, they suggest a model in which the autoregressive response is allotted one period in which to take effect, $C_t = \eta WC_{t-1}$. By contrast, other researchers do not seem to have problems with the idea that C_t in one spatial unit is regressed on C_t of other spatial units depending on a spatial weight matrix W , $C_t = \delta WC_t$.⁷ For that reason they do not preclude this specification in advance and suggest to determine whether the data can help to determine the most appropriate model. The specific-to-general approach is one way to test for this; estimate the simpler time-space recursive model and test whether the residuals of this model are free of any additional cross-sectional dependence. For this purpose, the Pesaran (2015) CD test may be used, whose null hypothesis is that the residuals are only weakly cross-sectionally dependent. This test

⁷Data frequency may also matter (daily, monthly, quarterly or annual data). Here we focus on annual data.

is based on the correlation coefficients between the time-series observations of each pair of spatial units with respect to a particular variable, in this case the residuals, resulting in $N \times (N - 1)$ correlations. Denoting these estimated correlation coefficients between the time-series for units i and j as $\hat{\tau}_{ij}$, the Pesaran (2015a) CD test is defined as $CD = \sqrt{(2T(N(N - 1)))} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \hat{\tau}_{ij}$. This two-sided test statistic has the limiting $N(0, 1)$ distribution as N and T go to infinity, and thus -1.96 and 1.96 as critical values at the 5% significance level. One advantage of this test is that it is not based on any (arbitrary) specification of the spatial weight matrix W . If the null is rejected, common factors have not adequately been accounted for. If the null is not rejected, it might be that WC_t is still relevant. The general-to-specific approach is the reverse way to test for this; estimate the model in Equation (1) (with or without common factors) and test the hypothesis whether $\delta = 0$.⁸

2.4 The quest for stationarity

A crucial issue is whether the dependent variable is stationary or alternatively is non-stationary, in which case it does have a unit root. The outcome of this test is determined by the regression model that is used to test for a unit root. The literature on unit roots first focused on time-series data, then on panel data, and finally on cross-sectionally augmented panel unit root tests (see Pesaran, 2015b, Ch. 31 for an overview). This second generation of panel unit root tests consists of PANIC (panel analysis of nonstationarity in idiosyncratic and common components) and CADF (cross-sectionally augmented Dickey-Fuller) tests, which take almost the same form as Equation 1. We present the differences compared to the CADF regression in Pesaran et al. (2013, eq.[19]) and then discuss the implications of these differences for our analysis.

First, spatial and spatiotemporal lags in the dependent variable, respectively WC_t and WC_{t-1} , are not accounted for ($\delta = \eta = 0$). Consequently, Pesaran et al.'s (2013) CADF test focuses on whether τ is equal to or smaller than 1. To test for this, the model is reformulated as $\Delta C_t = (\tau - 1)C_{t-1} + \dots = \tau' C_{t-1} + \dots$, after which it can be tested whether τ' is significantly smaller than 0 (stationarity) or equal to 0 (unit root). Note that in line with this reformulation of the model, \bar{C}_t is replaced by $\Delta \bar{C}_t$ and \bar{X}_{kt} by $\Delta \bar{X}_{kt}$, which helps to deal with some remaining serial correlation in the residuals. Second, instead of common coefficients across all units, coefficients are assumed to be unit-specific. By estimating the coefficients of the N unit-specific regressions by OLS, a series of N t-values is obtained for τ' , one for each spatial unit. Next, it is investigated whether the average t-value, which is negative since τ' is expected to be negative, is below or above a certain critical value, respectively pointing to stationarity or a unit root. These critical values are usually smaller than the common value of -1.96 at the 5% significance level and depend on the sizes of N and T , the number of time lags considered (in this study limited to 1), and the number of independent

⁸Since there is discussion in the spatial econometrics literature whether the specific-to-general or the general-to-specific approach is better, see Florax et al. (2003) and Mur and Angula (2009), we use a mix of both approaches.

variables in the model. Generally, these critical values fluctuate between -2 and -3. In Pesaran et al. (2013) it is assumed that the common factors are stationary and thus do not have a unit root. However, if the dependent variable, i.e. the consumption of cigarettes in Italy's provinces over the observation period, is nonstationary, so may be the cross-sectional averages of cigarette consumption over time. Fortunately, Kapetanios et al. (2011) show that the extension of cross-sectionally augmented econometric models to common factors that also have unit roots, although not straightforward, lead to similar results.

One implication of these differences is that due to the inclusion of WC_t and WC_{t-1} ($\delta, \eta \neq 0$) in Equation 1, we not only need to test whether $\tau' < 0$ ($\tau < 1$), but also whether $\delta + \tau + \eta - 1 < 0$ ($\delta + \tau + \eta < 1$). Yu et al. (2012) and Elhorst (2014, Section 4.6) have demonstrated that the model set out in Equation 1 becomes non-stable when the latter condition is not satisfied, or when the restriction $\delta + \tau + \eta = 1$ cannot be rejected, even if $\delta + \tau + \eta < 1$. One notable example of overlooking this condition is Fogli and Veldkamp (2011). In Table II of their study, they find that the response coefficient τ of the temporal lag C_{t-1} is 0.916 and η of the spatiotemporal lag WC_{t-1} is 0.570. Consequently, the sum of these two coefficients is greater than 1, i.e., the stationarity condition requiring that $\tau + \eta < 1$ (note $\delta = 0$ in this study) is not satisfied, pointing to potential misspecification problems that have not been identified. Unfortunately, there are no studies yet that have considered the estimation of Equation 1 with unit-specific coefficients for all variables. The most recent study of Aquaro et al. (2015) considers the extension of spatial autoregressive models to the case where the spatial coefficients differ across the spatial units, but these models are static rather than dynamic. Consequently, an overview of critical values of the t -statistic of $\delta + \tau + \eta - 1 < 0$ for different values of N and T is not available. In this study we will therefore report the average t -statistic of Pesaran et al.'s (2013) CADF test obtained from unit-specific regressions excluding WC_t and WC_{t-1} variables, and the t -statistic of $\delta + \tau + \eta - 1 < 0$ obtained when estimating Equation 1.

In line with Equation 1 and the regression set out in Pesaran et al. (2013, eq.[19]) underlying their CADF test, one may consider two cross-sectional averages for the dependent variable C , one at time t and one at time $t - 1$, and one for every independent variable X_k at time t . However, not all these common factors are necessary to obtain stationarity. As Pesaran et al. (2013, section 4.1) indicate, the number of common factors also depends on the number of independent variables taken up in the model and the exponent of the cross-section dependence, denoted by α . This exponent provides a characterisation of the degree of cross-sectional dependence in terms of the rate at which the average pair-wise correlation coefficient, also used to determine the CD-test, measured over all N units varies with N if N goes to infinity (Bailey et al., 2016).⁹ They find that $\bar{\rho}_N = O(N^{2\alpha-2})$. This implies that for values of α in the range of $[0, 1/2)$, the average correlation coefficient tends to go to zero very fast, pointing

⁹For detailed formulas we refer to their paper. Gauss code to calculate the α -test are made available in an online appendix to their paper.

to local spatial dependence comparable to a spatial arrangement of the units in the sample reflected by a binary contiguity matrix. Conversely, $\alpha = 1$ points to strong cross-sectional dependence and thus to the need to control for common factors. Values in between indicate moderate to strong cross-sectional dependence, where $\alpha = 3/4$ reflects a turning point. For values in the range of $[1/2, 3/4)$, the average correlation coefficient tends to go to zero slower than N but faster than \sqrt{N} , pointing to local spatial dependence comparable to a spatial arrangement of the units in the sample reflected by an inverse distance matrix (see Lee, 2002 for a more detailed explanation of this \sqrt{N} condition). For values of α in the range of $[3/4, 1)$, the average correlation coefficient tends to go to zero so slow, i.e. slower than \sqrt{N} , that each unit may say to affect all other units. This still points to the presence of common factors.

In sum, the quest for stationarity consists of two steps, each consisting of two tests. First, the α -test as well as the CD-test are used to see whether both weak and strong cross-sectional dependence have been effectively covered by the model, even if the number of common factors is limited. In this paper we consider \bar{C}_t , \bar{C}_{t-1} , and \bar{C}_t in combination with \bar{C}_{t-1} and \bar{X}_{kt} . The tests are applied to both the raw data and the residuals of the model. Second, we use the CADF-test and the stability test based on the restriction $\delta + \tau + \eta - 1 < 0$ to see whether the observed cigarette consumption patterns over space and time are stationary, depending on the common factors that have been taken up in the model.

3 Data source and descriptive statistics

3.1 Data source

The data used in this paper have been collected from the annual budget reports of the companies entrusted over time to manage the tobacco sector (*Regia Cointeressata dei Tabacchi* during the franchise period of 1869-1883, and the fully public *Azienda dei Tabacchi* during 1884-1913).¹⁰ The public monopoly of tobacco was established in Italy since 1862, immediately after the political unification of 1861. However, Venetia and Latium were annexed to Italy only in, respectively, 1866 and 1870, as a result of which we do not have systematic data for the eight provinces of Venetia and for Latium before 1870. In addition, Sicily joined the Italian public monopoly only in 1877, as a result of which systematic data for the seven Sicilian provinces for the years before 1877 are not available either. This explains why our investigation period covers the years from 1877 to 1913, for a total of 37 years. The cross sectional units are the 69 Italian provinces of the time. As detailed in Ciccarelli (2012), the historical sources on tobacco in Italy refer to legal sales from local tobacco warehouses, distributed uniformly across Italy's territory. Since the data are reported in both physical (kilograms) and money terms (Italian lire), data on cigarette "consumption" fully reflect legal sales of cigarettes at the provincial level. In addition, it is

¹⁰See Madson (1916, pp. 105-107) for details on the institutional settings.

important to note that the price of cigarettes was fixed by the State at the national level, and that this variable is obtained by dividing, province by province and year by year, the sales data in monetary terms (expenditure on cigarettes in lire) by the sales data in physical terms (sales of cigarettes in kilograms).

Price differentials across provinces in a particular year, even though tobacco prices were set at the national level, are due to quality differences, but generally these differences are small. Data only available at the national level show in particular that just four types of cigarettes (each of low quality) covered more than 80 percent of the market, with a peak value above 95 percent in 1899.¹¹ We converted the nominal price of cigarettes in real terms using the cost of living index reported in Fenoaltea (2011, p. 128). In addition, it should be noted that for the historical period considered here, annual estimated values of income at the regional or provincial level (respectively NUTS 2 and 3 units) are not available. This is typical for historical studies where usually only benchmarks estimates (usually decennial) of regional GDP are available, starting from the second half of the nineteenth century. Therefore, in the present study we use as a proxy for annual provincial income during 1877-1913 the same variable, based on fiscal data, as is used in Ciccarelli and De Fraja (2014). Provincial income data were also converted in real terms using the the cost of living index reported in Fenoaltea (2011, p. 128). Finally, population and literacy data at the provincial level for the years 1871, 1881, 1901, 1911 are from the population censuses, while those for the remaining years were obtained by linear interpolation of provincial figures.¹²

3.2 Spatio-temporal patterns

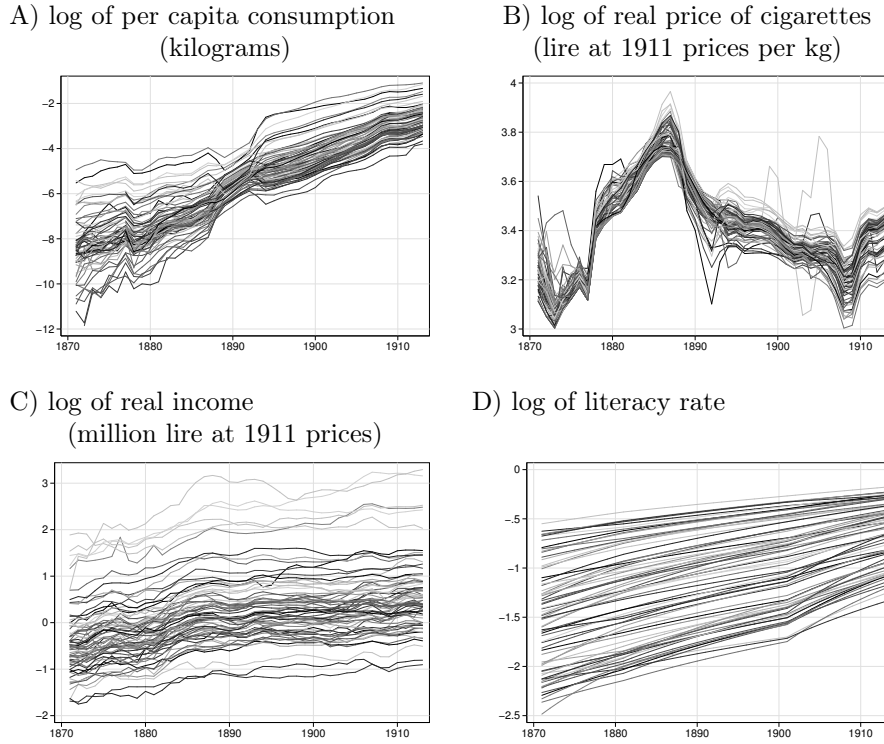
This section illustrates the evolution over time (1871-1913) and space (69 Italian provinces) of the main variables included in the proposed dynamic model for panel data of cigarette diffusion.

Panels A, B, C, and D of Figure 2 graph the temporal evolution of per capita consumption of cigarettes in physical terms, the real price of cigarettes, real income, and literacy rates of all 69 provinces in the sample. Except for the price of cigarettes, each panel shows a clear rising trend. Panel A combines the temporal pattern of per capita cigarette consumption at the national level illustrated in Figure 1 and the spatial pattern across these provinces in different years illustrated in Figure 3 (to be discussed shortly). It emerges that consumption had the tendency to move together in the various provinces, but that the degree of co-movement differs from one province to another. Turning to real in-

¹¹The four types of cigarettes were the “spagnolette of 3rd quality” made of foreign tobacco; the “spagnolette of 3rd quality” made of national tobacco including both the so called “macedonia” and the so called “virginia and maryland”; the “spagnolette of 4th quality” made of national tobacco and usually called “nazionali”. The cigarettes “nazionali” were introduced in 1890, the “virginia and maryland” cigarettes were introduced in 1897. In 1895, the manufacturing of “spagnolette of 3rd quality” made of foreign tobacco was instead completely abandoned.

¹²In 1891 the population census was not taken. See Ciccarelli and Weisdorf (2016) for a long-term analysis of the historical evolution of literacy in Italian provinces.

Figure 2. Consumption and price of cigarettes, income, and literacy rates: provincial trends 1871-1913^a



^a Panels A and B: data for Sicilian provinces for the period 1871-1876 are not available.

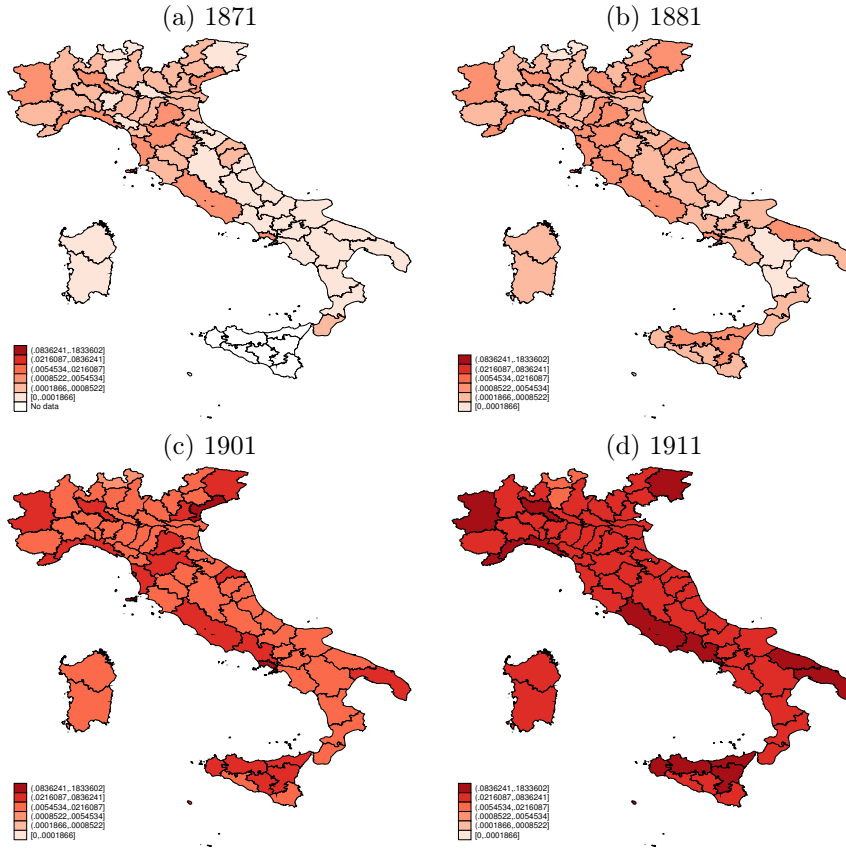
come (illustrated in panel C), basic calculations show that it more than tripled in the long run, and that after a period of growth during 1870-1890, the 1890s marked a period of stagnation. However, over the period 1900-1913 income increased at a particularly fast rate, especially so in selected provinces. The rising trend of literacy rates (panel D) is tied to the gradual diffusion of a free and mandatory primary public school system. The price of cigarettes, graphed in panel B, increased by law with the reforms of 1878, 1885 and 1909-1910. The price reform of 1878 was particularly severe and produced a reduction of cigarettes sales of some 30 percent relative to 1877, when expressed in kg, and of some four percent in money terms. However, the 1878 price reform was an isolated case, and can essentially be interpreted as a first test of the elasticity of demand for cigarettes run by policy makers, when the market was still of limited size.¹³ Since the mid-1880s the rising trend reverted, and the price of

¹³This is clearly evident in the parliamentary proceedings of 1878, reporting tables of State's

cigarettes reduced considerably, partly due to the replacement of hand-rolled by machine-rolled cigarettes, which led to a considerable reduction in production costs, and ultimately in the price of cigarettes.

Figure 3 illustrates, on a common scale based on the percentiles of the pooled dataset, the spatial distribution of cigarettes consumption among Italian provinces in selected years: 1871, 1881, 1901, and 1911. The maps show that the territorial diffusion of cigarettes followed a clear North-South gradient. In 1871

Figure 3. Per capita consumption of cigarettes, selected years (kg)^a



^a Data for Sicilian provinces are only available starting 1877. Class intervals based on the percentiles (pc1; pc5; pc25; pc50; pc75; pc95; pc99) of the whole distribution of the pooled data (all provinces, all years).

and 1881 cigarettes consumption in the South was negligible or relatively low, as indicated by the prevailing light colours. At the beginning of the century the

expected profits for the years 1879-1883 induced by the price reform of 1878 (Camera dei Deputati, 1878, pp. 13-19).

distribution was instead rather uniform, confirming the rapid territorial spread of cigarettes consumption. It is also interesting to note that that consumption in 1911 was typically high in the provinces with the main city capitals, reinforcing a trend partly visible also in the previous years. Dark colours appear in particular for the provinces of Turin, Milan, and Genoa in the North-West; Venice in the North-East (at least up to 1901); Rome in the center, and Naples and Palermo in the South.¹⁴

Overall, we observe in the data the same three stylized facts Halleck-Vega and Elhorst (2016) found for unemployment rates in the Netherlands across 12 provinces over the period 1973-2013: (i) Cigarette consumption tends to be strongly correlated over time. Based on Figure 1 and Figure 2 panel A, the level of consumption in period $t - 1$ is a good predictor of the level of consumption in period t ; (ii) the level of consumption at the provincial level parallels its counterpart at the national level. Based on Figure 2 panel A, the level of consumption at the national level is a good predictor of the consumption at the provincial level; (iii) the levels of consumption at the provincial level are locally correlated across space. Based on the local clustering of provinces in Figure 3, the levels of consumption in the neighbors of a particular province are a good predictor of the level of consumption in that province itself. The main difference with Halleck-Vega and Elhorst (2016) is that the data in this study are anything but stationary, one of the main characteristics of diffusion processes. Whether the model proposed in Equation (1) can adequately describe these data is the topic of the next section.

4 Empirical findings

We first test for cross-sectional dependence using Pesaran's (2015a) CD-test based on the consumption patterns of the 69 provinces over the period 1877-1913 ($N=69$, $T=37$). The result is 190.42 with an average pairwise correlation coefficient of 0.89. This outcome is highly statistically significant, indicating that cross-sectional dependence needs to be accounted for. The exponent α -test of Bailey et al. (2015) amounts to $\alpha = 1.002$ with standard error 0.018, which points to the existence of strong cross-sectional dependence, and thus the need to control for common factors. When testing for stationarity using the CADF test (for the moment without common factors), we find a t-value of the negative outcome for τ' of only -0.159, indicating that the hypothesis of stationarity needs to be rejected in favor of a unit root.

Table A1 in the appendix reports the correlation matrix between the independent variables and their spatially lagged counterparts. Generally, these variables do not highly correlate with each other, except for the price of cigarettes

¹⁴In June 1911, when the population census was taken, the population in Italy amounted to 34,671,377 individuals. The provincial percentage shares of the five most populated provinces were as follows: Milan (4.98 %), Naples (3.78 %), Rome (3.76 %), Turin (3.50 %), and Genoa (3.03 %). The average provincial population was 502,484, while the minimum was 129,928 (Sondrio), and the maximum was 1,726,548 (Milan).

in the own province with that in neighboring provinces (0.977), and the level of education in the own province with that in neighboring provinces (0.934). The first correlation is explained by the fact that prices are set by the State at the national level. The reason that the coefficient is close to but not equal to 1 is due to minor quality differentials (see Section 3). The second correlation is due to the limited information that is available to construct data series on education. At four points in time (1871, 1881, 1901, and 1911, corresponding to the census years) we observe variation across space, but this is apparently insufficient to discriminate between the impact of education caused by the own region and that by neighboring regions. Hence, both the price of cigarettes and the level of education observed in neighboring provinces have been left aside in this study.

The last column of Table 1 reports the results of the model specification proposed in Equation 1 and the first five columns the results of five simpler model specifications to test the sensitivity of the proposed model to potential alternatives. Step by step we will show why these simpler models, which have been considered in many previous studies before, fall short and why the proposed model specification suffices. Models 1 and 2 embody the dynamic spatial panel data model without common factors ($\Gamma = \Pi = 0$) but with time-period fixed effects, representing the common approach characterizing the spatial econometrics literature. The specification of the spatial weight matrix W in model 1 takes the form of a binary contiguity (BC) matrix and in model 2 of an inverse distance (ID) matrix. In both cases, non-stationarity appears to be a problem.

On adopting the ID matrix, the sum of the serial, spatial, and spatiotemporal autoregressive coefficients, $\tau + \delta + \eta$, of the variables C_{t-1} , WC_t , and WC_{t-1} turns out to be 1.100. Consequently, $\tau + \delta + \eta - 1$ is greater rather than smaller than 0 and the corresponding t-value of 0.805 positive rather than negative. This result pointing to non-stationarity does not really come as a surprise. Most studies adopt a sparse binary contiguity matrix rather than a dense inverse distance matrix when modeling diffusion processes (see Section 2.1). In an overview study of 29 marketing studies using spatial econometric methods, Elhorst (2017) finds that a vast majority adopts the BC matrix, especially when focusing on diffusion processes. This is perfectly in line with the proposed spatial econometric specification. The BC matrix specifies which units are neighbors of the leading regions where new ideas or habits are born. The provinces where per capita cigarette consumption started to rise first and remained (almost) constantly above twice the national average are Leghorn, Naples and Venice. The coefficients δ and η in Equation 2 on their turn measure the propagation of this habit from one province to another, i.e., from the leading region to their neighbors, to the neighbors of these neighbors, and so on.

On adopting the BC matrix (model 1), the sum of the three coefficients is smaller than 1 (0.985), but the t-value of -1.555 regarding $\tau + \delta + \eta - 1$ indicates that the hypothesis of stationarity still needs to be rejected in favor of a unit root. The explanation for these findings is that the controls for time period effects are not sufficient to adequately cover the non-stationarities caused by the diffusion process. Halleck-Vega and Elhorst (2016) demonstrate that time-

Table 1. Estimation results^a

	M1	M2	M3	M4	M5	M6
C_{t-1} (τ)	0.912 (101.03)	0.880 (101.33)	0.861 (47.87)	0.967 (44.75)	0.811 39.44	0.828 (61.31)
WC_t (δ)	0.435 (16.08)	0.672 (26.29)	-0.230 (-8.61)	-0.234 (-8.70)	-0.242 (-9.11)	-0.236 (-8.64)
WC_{t-1} (η)	-0.362 (-12.67)	-0.452 (-9.88)	-0.230 (-7.69)	0.144 (3.54)	-0.265 (-8.64)	0.183 (6.06)
Income (β_2)	-0.074 (-2.78)	-0.051 (-1.93)	-0.081 (-1.68)	0.020 (0.35)	-0.026 (-0.45)	0.019 (0.50)
Price (β_1)	-0.124 (-2.32)	-0.154 (-2.67)	0.345 (6.81)	-0.225 (-3.52)	0.157 (1.62)	-0.088 (-2.28)
Education (β_3)	-0.062 (-0.41)	0.034 (0.20)	-1.210 (-5.56)	2.012 (7.57)	-5.930 (-6.18)	0.170 (1.05)
$W \times$ Income (θ_1)	0.140 (3.81)	0.372 (3.52)	0.041 (0.63)	0.809 (10.42)	0.542 (5.53)	0.066 (1.35)
R^2 corrected	0.857	0.840	0.993	0.991	0.994	0.996
Spatial effects (μ)	YES	YES	YES	YES	YES	YES
Time dummies (ξ)	YES	YES	NO	NO	NO	NO
Common factors (Γ, Π)	NO	NO	\bar{C}_t	\bar{C}_{t-1}	\bar{C}_t, \bar{X}_t	\bar{C}_t, \bar{C}_{t-1}
CADF-test (t-value)	-2.388	-2.388	-1.507	-0.164	-1.836	-2.596
$\hat{\tau} + \hat{\delta} + \hat{\eta} - 1$	-0.015 (-1.55)	0.101 (0.81)	-0.598 (-27.87)	-0.125 (-4.55)	-0.697 (-30.66)	-0.225 (-30.32)
CD test (residuals)	-0.407	-3.098	65.807	162.160	55.066	-1.894
α -test	0.643	0.674	0.877	0.542	0.896	0.591
(standard error)	(0.042)	(0.046)	(0.093)	(0.055)	(0.026)	(0.100)

^a Bias-corrected ML estimates (Yu *et al.*, 2008). Common factors and spatial fixed effects are not reported for reason of space. t-values in parenthesis. Models 1, and 3-5 (M1, M3-M6) use a normalized binary contiguity W matrix, model 2 (M2) a normalized inverse distance W matrix. When evaluating the sum $\hat{\tau} + \hat{\delta} + \hat{\eta} - 1$, numbers need not add up due to rounding.

period fixed effects are a special case of common factors with $\Gamma_1 = \gamma \iota_N$ and $\Gamma_2 = 0$, i.e. the first common factor at time t with unit-specific coefficients is replaced by a time dummy with a common coefficient γ for all regions and the second common factor is set to zero. Apparently, this simplification is too restrictive to obtain a stable model. This also follows from the results obtained for the R^2 , which for models 1 and 2 is substantially smaller than for the models including common factors to be discussed below.¹⁵

The CD-test applied to the residuals of models 1 and 2 are respectively -0.407 for the BC matrix and -3.016 for the ID matrix. Although the CD

¹⁵This R^2 does not include the contribution of WC_t ; instead it determines the explanatory power of the reduced form equation obtained by moving δWC_t from the right to the left hand side of Equation 1 and multiplying both sides by the spatial multiplier matrix $(I_N - \delta W)^{-1}$. See Elhorst (2014, section 3.5.1 for details).

has not been developed as a device for selecting the right W matrix, it can nonetheless be interpreted as another result demonstrating that the BC matrix outperforms the ID matrix. When the CD matrix is adopted and spatial lags are accounted for in the dependent variable and the independent variables, we no longer find evidence in favor of any additional cross-sectional dependence. By contrast, when the ID matrix is adopted, not all cross-sectional dependence appears to be covered. In addition, the R^2 of model 1 based on the BC matrix is greater than that of model 2 based on the ID matrix. We therefore no longer consider the ID matrix, but only the BC matrix.

Model 3 represents the dynamic spatial panel data model with controls for common factors of the dependent variable at time t only. It generalizes model 1 in that the common factor Γ_1 has unit-specific coefficients rather than a common coefficient for all units. For this reason we do not longer control for time dummies (see Table 1) since their impact is covered by the common factor. Taking them up both would cause perfect multicollinearity. One reason to control for one common factor at time t only rather than at both time t and time $t-1$ might be to reduce the number of parameters to be estimated. The model already contains N unit specific-intercepts μ , each additional common factor implies another set of N parameters to be estimated. To avoid that this number grows too large, more common factors at time $t-1$ (or at time t , see below) may be left aside. The results show that this is an extremely effective way to obtain a stable model, the sum of the coefficients τ , δ and η falls to 0.402, which is significantly smaller than 1, but the downside is that the CD-test applied to the residuals of this model takes a significant value of 65.807, and the α -test a value of 0.877 which is still in the range of $[3/4, 1)$. These results indicate that this model does not account for both weak and strong cross-sectional dependence effectively.

A similar result is obtained when controlling for a common factor at time $t-1$ only, represented by model 4. With t-values of only -1.507 and -0.164 the CADF-tests point out that the stationarity condition is not satisfied when either \bar{C}_t or \bar{C}_{t-1} is controlled for, while the CD-test statistic with a value of 162.16 if \bar{C}_t is left aside still points to additional strong cross-sectional dependence among the residuals. The same happens to model 5, i.e. when cross-sectional averages of the dependent and the independent variables are controlled for at time t only. The parameter estimates and the test statistics are very similar to those of model 3, indicating that this model, just as model 3, does not account for both weak and strong cross-sectional dependence effectively. Only in model 6, when common factors of the dependent variable at both points in time are controlled for, we obtain an insignificant CD-test statistic of -1.894 and an α -test of 0.591 below the turning point of $3/4$ indicating that both weak and strong cross-sectional dependence have been adequately tackled, as well as a t-value of -2.596 for the CADF test and of -12.555 regarding the stationary condition $\delta + \tau + \eta - 1 < 0$ indicating that the hypothesis of stationarity no longer needs to be rejected in favor of unit root. Interestingly, these results are in line with the approach developed by Halleck-Vega and Elhorst (2016) simultaneously accounting for serial dynamics and both weak and strong cross-sectional dependence within

one framework, generalizing the two-stage estimation procedure pioneered by Bailey et al. (2016). According to this approach cross-sectional averages of the dependent variable both at time t and $t - 1$ need to be controlled for; not doing so leads to a misspecified model.

Another way to look at cross-sectional dependence and stationarity is to consider the coefficient estimates of the different models and the direct and spillover effects derived from them. Not effectively accounting for cross-sectional dependence and stationarity may lead to spurious and thus counterintuitive results, as we will show below.

Korniotos (2010) interprets the coefficients of the temporal and space-time lags of the dependent variable (C_{t-1} and WC_{t-1}) as measures of what he calls the relative strength of internal and external habit persistence, where external habit persistence reflects the information citizens of a particular region pick up from their neighbors, in this case an upcoming trend to smoke cigarettes. Both the internal and external habit persistence parameters are therefore expected to be positive, which is the case for models 4 and 6, but not for models 1, 2, 3, and 5. Table 2 reports the estimated short-term and long-term effects derived from the coefficient estimates of model 6 using Equation 2. The results are plausible since they are in line with standard economic theory.¹⁶ The total

Table 2. Short-term and long-term direct and spillover effects^a

MODEL 6: Common factors \bar{C}_t and \bar{C}_{t-1} included						
	short-term			long-term		
	direct	indirect	total	direct	indirect	total
income	0.015 (0.37)	0.055 (1.16)	0.070 (2.24)	0.076 (0.31)	0.309 (1.07)	0.385 (2.20)
price	-0.089 (-2.38)	0.018 (2.30)	-0.071 (-2.38)	-0.537 (-2.25)	0.145 (1.49)	-0.392 (-2.32)
educ	0.176 (1.06)	-0.035 (-1.05)	0.141 (1.06)	1.056 (1.05)	-0.280 (-0.87)	0.775 (1.05)

^a Model based on binary contiguity matrix. t-values in parenthesis.

income elasticity in the short-term amounts to 0.070 and in the long-term to 0.385. If income increases people consume more cigarettes, like a normal good. Both numbers are also statistically significant. The own effect and the spillover effect of income are also positive, though insignificant. The total price elasticity in the short-term is -0.071 and in the long-term it is -0.392. If prices increase, people reduce their consumption, as expected. In contrast to income, not only the total effect is significant, but also the direct effect and the spillover effect. If the price of cigarettes rises in the own province, consumption falls by 0.089 in the short-term and by 0.537 in the long-term. However, this fall is compensated by a positive spillover effect of 0.018 in the short-term and 0.145 in the long-

¹⁶For the sake of completeness, the results for the other models are reported in Table A2 of the Appendix.

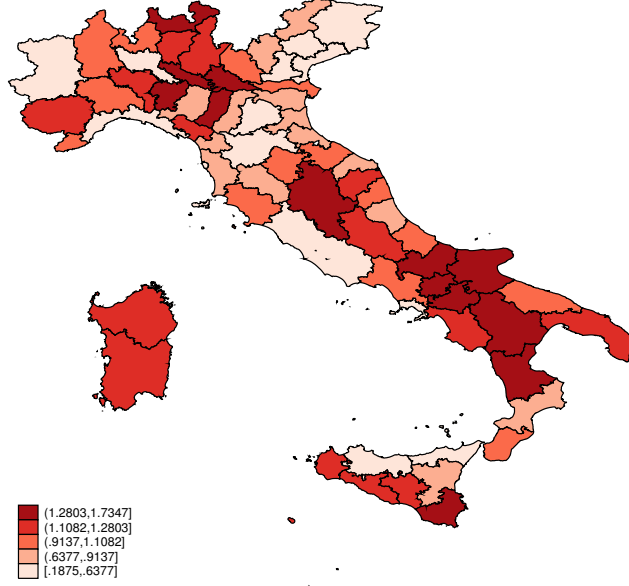
term. This spillover effect can be interpreted as a substitution effect and is in line with studies that investigated cigarette consumption in 46 U.S. States over the period 1963-1992 discussed in Section 2.1. Since prices are set by the State in principle, this substitution effect should be read here as quality rather than price substitution. In sum, it may be concluded that cigarettes have been experienced as a normal consumption good at the time of their introduction in Italy.

The proposition that the better educated, here measured by the literacy rate, were the first to start smoking, suggested by Panels A and D of Figure 2, is neither rejected nor corroborated by the effects estimates. Although the short-term and long-term direct and total effects are positive, neither of these effects appears to be significant. The explanation might be that the relation between education and cigarettes consumption is driven by both positive and negative effects. The relation may be expected to be positive since cigarettes represented a new trendy good at the time of their introduction, thereby meeting the social needs of educated people, but negative on the ground that the better educated people were aware of the negative consequences of smoking to their health. For example, the Italian Parliament considered approving a law banning cigarettes smoking to the young in 1907 (Ciccarelli and De Fraja, 2014, p. 167). However, apart from the fact that the law was not approved, it seems safer to claim that in 19th century there was no generalized public awareness yet of the negative consequences of smoking for health.

Importantly, these plausible results could only be achieved by controlling for common factors at both time t and $t - 1$. If common factors are not controlled for (model 1, Table A2), we would conclude that the short-term direct effect of income is negative and significant, and that all the short-term effects and the long-term total effect of education although not significant are also negative. If common factors are controlled for but only at time t (model 3 or model 5), we would conclude that the short-term and long-term total effects of education are not only negative but also significant. These results are counterintuitive. Finally, if common factors are controlled for but only at time $t - 1$ (model 4), we would indeed conclude that the impact of education is strongly positive and significant. On the other hand, the opposite signs and the relative large magnitudes of the direct and spillover effects found for education, as well as income, in model 4 point to misspecification problems, as have been identified earlier by the CADF-test and the CD-test on the residuals of this model.

Finally, we have estimated the extent to which each province responds to the national trend. These parameters, which are region-specific and denoted by γ , can be found either by dividing the elements of Γ_1 by $1 - \delta$ or by dividing the elements of Γ_2 by $-\tau - \eta$. Halleck-Vega and Elhorst (2016) provide the details behind these calculations and also show that these calculations produce significantly different results. In this paper we used the latter method, since it is based on the relative strength of both internal and external habit persistence (Korniotis, 2010), which were both found to be positive and significant. Studies controlling for common factors have the tendency not to report the γ parameter estimates, probably because of their large number. This is a pity since they

Figure 4. The diffusion of cigarettes in Italy, 1877-1913: trendsetters and trend followers ^a



^a The map reports the 69 coefficients of the diagonal matrix Γ_2 , scaled by $(-\tau - \eta)$. Numbers less than 1 denote trendsetters and are represented in the map by light colours, while those greater than 1 denote trend followers and are represented in the map by dark colours.

throw more light on the problem at hand. In line with Figure 3, provinces turn out to be trendsetters if $\gamma < 1$, and trend followers if $\gamma > 1$. As illustrated by Figure 4, the general picture we find is that provinces with major urban centers like Palermo and Naples in the South, Rome and Florence in the Center, and Genoa, Turin, Milan, and Venice in the North are trendsetters. Low density rural provinces located in the Central and Southern Apennines, but also those located along the Po Valley, are instead typically trend followers.

5 Conclusion

To model diffusion processes characterized by non-stationary data, a dynamic spatial panel data model with common factors is proposed. This model is used to provide a better understanding of the propagation over space and time of cigarette demand in 69 Italian provinces over the period 1877-1913. The rapid diffusion of cigarettes was sustained by both supply and demand factors. On the one hand, technological progress induced by the replacement of hand-rolled by machine-rolled cigarettes implied a considerable reduction in production costs,

and ultimately in the price of cigarettes. On the other hand, the rapid increase in real income, especially pronounced after the turn of the century also contributed to sustaining the demand for cigarettes. The observations that cigarettes represented a new and trendy product at the turn of the 19th century in Italy and that the data over this period are not stationary make the modeling of its diffusion process an interesting and challenging topic of research.

The Pesaran (2015a) CD-test and the exponent α -test of Bailey et al. (2015) show that both weak and strong cross-sectional dependence are important drivers of the propagation of cigarette demand over space and time and thus should be accounted for. If common factors are accounted for but only at one particular point in time rather than both t and $t - 1$, the Pesaran CD-test applied to the residuals of the model points to additional cross-sectional dependence, indicating that weak and strong cross-sectional dependence and the mutual relationship between them using this setup are still not modeled adequately. Only when common factors at both points in time are controlled for, we obtain a model whose residuals are free of any additional cross-sectional dependence, and a model that is stationary in the sense that both the condition $\tau + \delta + \eta - 1$ in the dynamic spatial panel data model and $\tau' = \tau - 1$ in Pesaran et al.'s unit-specific regressions of this model (excluding the variables WC_t and WC_{t-1}) are significantly smaller than 0.

The estimation results obtained when using the proposed model show that cigarettes are a normal good. The income elasticity amounts to a significant value of 0.385 in the long-term and the price elasticity to a significant value of -0.392 in the long term. The long-term direct effect of a price increase in a region itself is 0.537, while the spillover effect to other regions is 0.145. The latter two elasticities are also significant. The positive spillover effect is in line with other spatial econometric studies which investigated cigarette demand in the U.S. states over a more recent period. Although both cigarette consumption and the level of education measured by the literacy rate increased considerably over the observation period, the relationship between them, although positive, is not found to be significant.

Appendix

Table A1. Correlation matrix^a

	price	income	educ	W*price	W*income	W*educ
price	1.000					
income	0.116	1.000				
education	-0.192	0.330	1.000			
W*price	0.977	0.026	-0.220	1.000		
W*income	0.047	0.223	0.258	0.084	1.000	
W*educ	-0.244	0.156	0.934	-0.246	0.345	1.000

^a W = binary contiguity matrix.

Table A2. Short-term and long-term direct and spillover effects^a

MODEL 1: No common factors, but time dummies included						
	short-term			long-term		
	direct	indirect	total	direct	indirect	total
income	-0.059	0.175	0.116	5.845	-4.163	1.683
	(-2.31)	(3.40)	(2.11)	(0.03)	(-0.02)	(0.02)
price	-0.132	-0.088	-0.220	-8.560	5.899	-2.660
	(-2.34)	(-2.28)	(-2.33)	(-0.04)	(0.02)	(-0.02)
educ	-0.058	-0.038	-0.096	11.690	-13.338	-1.647
	(-0.38)	(-0.37)	(-0.37)	(0.03)	(-0.03)	(-0.02)

MODEL 3: Common factor \bar{C}_t included						
	short-term			long-term		
	direct	indirect	total	direct	indirect	total
income	-0.082	0.049	-0.034	0.611	-0.681	-0.070
	(-1.67)	(0.82)	(-0.84)	(0.04)	(-0.04)	(-0.84)
price	0.352	-0.070	0.282	-1.715	2.296	0.581
	(7.17)	(-5.44)	(7.26)	(-0.02)	(0.03)	(6.85)
educ	-1.214	0.242	-0.972	4.845	-6.848	-2.003
	(-5.50)	(4.64)	(-5.52)	(0.02)	(-0.03)	(-5.31)

MODEL 4: Common factor \bar{C}_{t-1} included						
	short-term			long-term		
	direct	indirect	total	direct	indirect	total
income	-0.029	0.704	0.676	-15.496	22.532	7.036
	(-0.48)	(9.24)	(13.27)	(-0.02)	(0.03)	(4.03)
price	-0.233	0.047	-0.186	-6.235	4.303	-1.932
	(-3.56)	(3.27)	(-3.54)	(-0.01)	(0.01)	(-2.72)
educ	2.052	-0.413	1.639	32.649	-15.596	17.053
	(7.61)	(-5.59)	(7.68)	(0.01)	(-0.00)	(3.75)

MODEL 5: Common factors \bar{C}_t and \bar{X}_{kt} included						
	short-term			long-term		
	direct	indirect	total	direct	indirect	total
income	-0.060	0.478	0.418	-0.257	1.001	0.745
	(-1.04)	(5.73)	(5.23)	(-0.01)	(0.03)	(5.23)
price	0.162	-0.033	0.128	0.318	-0.091	0.278
	(1.61)	(-1.58)	(1.61)	(0.10)	(-0.01)	(1.62)
educ	-6.044	1.254	-4.790	-29.596	21.063	-8.534
	(-6.29)	(5.11)	(-6.33)	(-0.04)	(0.03)	(-6.38)

^a This table reports the estimated short-term and long-term direct and spillover effects for models M1, M3, M4, and M5 for the sake of completeness. As explained in the main text, estimates from model M6 are preferable. All models are based on binary contiguity matrix. t-values in parenthesis. Results of Model 2 based on the inverse distance matrix are not reported due to instability found in Table 1. The results of Model 1 are reported for reasons of comparison, despite their instability.

References

- Anselin, L., Le Gallo, J., and Jayet, H., 2008, “Spatial panel econometrics”. In: Matyas, L., and Sevestre, P. (eds) *The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice*, pp. 901-969, 3rd edn. Kluwer: Dordrecht.
- Aquaro, M., Bailey, N., and Pesaran, M.H., 2015, “Quasi maximum likelihood estimation of spatial models with heterogeneous coefficients”, USC-INET Research Paper No. 15-17. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2623192>.
- Bailey, N., Holly, S., and Pesaran, M.H., 2016, “A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence”, *Journal of Applied Econometrics*, 31, pp. 249-280.
- Bailey, N., Kapetanios, G., and Pesaran, M.H., 2015, “Exponent of cross-sectional dependence: estimation and inference”. *Journal of Applied Econometrics*, doi: 10.1002/jae.2476.
- Baltagi, B.H., and Levin, D., 1986, “Estimating dynamic demand for cigarettes using panel data: the effects of bootlegging, taxation and advertising reconsidered”, *The Review of Economics and Statistics*, 68, pp. 148-155.
- Baltagi, B.H., and Levin, D., 1992, “Cigarette taxation: raising revenues and reducing consumption”, *Structural Change and Economic Dynamics*, 3, pp. 321-335.
- Baltagi, B.H., and Li, D., 2004, “Prediction in the panel data model with spatial autocorrelation”, In: Anselin, L., Florax, R., and Rey, S.J. (eds.), *Advances in Spatial Econometrics: Methodology, Tools, and Applications*, pp. 283-295, Springer: Berlin.
- Bollinger, B., and Gillingham, K., 2012, “Peer effects in the diffusion of solar photovoltaic panels”, *Marketing Science*, 31, pp. 900-912.
- Brady, R.R., 2011, “Measuring the diffusion of housing prices across space and time”. *Journal of Applied Econometrics*, 26, pp. 213-231.
- Burridge, P., 1981, “Testing for a common factor in a spatial autoregression model”, *Environment and Planning A*, 13, pp. 795-400.
- Camera dei deputati, 1878, Progetto di legge no. 38, 8 maggio 1878, Atti Parlamentari, Sessione del 1878, XII Legislatura.
- Cappellari della Colomba, G., 1866, *Le imposte di confine. I monopoli governativi e i dazi di consumo in Italia*, Stamperia Reale: Florence.
- Carrion-i-Silvestre, J.L., and Surdeanu, L., 2016, “Productivity, infrastructure and human capital in the Spanish regions, *Spatial Economic Analysis*, doi: 10.1080/17421772.2016.1189089.

- Chudik, A., M.H. Pesaran, and Tosetti, E., 2011, “Weak and strong cross-section dependence and estimation of large panels”, *The Econometrics Journal*, 14, pp. C45-C90.
- Ciccarelli, C., 2012, “The consumption of tobacco in Italy: national and regional estimates, 1871-1913”, *Rivista di storia economica*, 28, 3, pp. 409-452.
- Ciccarelli, C., and De Fraja, G., 2014, “The demand for tobacco in post-unification Italy”, *Chimetica*, 8, 2, pp. 145-171.
- Ciccarelli, C., Pierani, P. and S. Tiezzi (2014), “Secular trends in tobacco consumption: the case of Italy, 1871-2010”, Università degli Studi di Siena, WP n. 700.
- Ciccarelli, C., and Weisdorf, J., 2016, “The effect of the Italian unification on the comparative regional development in literacy, 1821-1911”, CEIS Tor Vergata Research Paper Series No. 392
- Cross, G.S., and Proctor, R.N., 2014, *Packaged Pleasures. How Technology and Marketing Revolutionized Desire*, the University of Chicago Press: Chicago and London.
- Debarsy, N., Ertur, C., and LeSage, J.P., 2012, “Interpreting dynamic space-time panel data models”, *Statistical Methodology*, 9, pp. 158–171.
- Diana G., 1998, “Le sigarette fatte a mano”, *La voce del tabaccaio*, 71, 20, pp. 44–47.
- Diana G., 1999, “La storia della sigaretta”, *Il Tabacco*, 7, pp. 25-32.
- Elhorst, J.P., 2014, *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, Springer-Verlag: Berlin Heidelberg.
- Elhorst, J.P., 2017, “Spatial models”. In: Leeflang, P.S.H, Wieringa, J.E., Bijmolt, T.H.A., and Pauwels, K.H. (eds.), *Advanced Methods for Modelling Markets*, Springer: Berlin, forthcoming, doi: 10.1007/978-3-319-53469-5_6.
- Elhorst, J.P., Piras, G., and Arbia, G., 2010, “Growth and convergence in a multi-regional model with space-time dynamics”, *Geographical Analysis*, 42, pp. 338-355.
- Elhorst, J.P., Zandberg, E., and Haan, J. de, 2013, “The impact of interaction effects among neighboring countries on financial liberalization and reform; A dynamic spatial panel data approach”, *Spatial Economic Analysis*, 8, pp. 293-313.
- Ertur, C., and Musolesi, A., 2016, “Weak and strong cross-sectional dependence: a panel data analysis of international technology diffusion”, *Journal of Applied Econometrics*, doi: 10.1002/jae.2538.

- Fenoaltea, S., 2011, *The Reinterpretation of Italian Economic History. From Unification to the Great War*, Cambridge University Press: Cambridge, NY.
- Florax, R.J.G.M., Folmer, H., and Rey, S.J., 2003, "Specification searches in spatial econometrics: the relevance of Hendry's methodology", *Regional Science and Urban Economics*, 33, pp. 557-579.
- Fogli, A., and Veldkamp, L., 2011, "Nature or nurture? Learning and the geography of female labor force participation", *Econometrica*, 79, pp. 1103-1138.
- Halleck-Vega, S., and Elhorst, J.P., 2015, "The SLX model", *Journal of Regional Science*, 55, pp. 339-363.
- Halleck-Vega, S.M., and Elhorst, J.P., 2016, "A regional unemployment model simultaneously accounting for serial dynamics and strong cross-sectional dependence", *Regional Science and Urban Economics*, 60, 85-95.
- Halleck-Vega, S.M., and Elhorst, J.P., 2017, "Regional labor force participation across the European Union: a time-space recursive modeling approach with endogenous regressors", *Spatial Economic Analysis*, 12, 422-439.
- Hutson, J.B., 1937, "Consumption and production of tobacco in Europe", *Technical Bulletin*, no. 587, United States Department of Agriculture: Washington, D.C.
- Kelejian, H.H., and Piras, G., 2014, "Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes", *Regional Science and Urban Economics*, 46, pp. 140-149.
- Kapetanios, G., Pesaran, H.M., and Yamagata, T., 2011, "Panel with non-stationary multifactor error structure", *Journal of Econometrics*, 160, pp. 326-348.
- Korniotis, G.M., 2010, "Estimating panel models with internal and external habit formation", *Journal of Business and Economic Statistics*, 28, pp. 145-158.
- Lee, L.-F., 2002, "Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models", *Econometric Theory*, 18, pp. 252-277.
- Lee, L.-F., and Yu, J., 2010a, "A spatial dynamic panel data model with both time and individual fixed effects", *Econometric Theory*, 26, pp. 564-597.
- Lee, L.-F., and Yu, J., 2010b, "Some recent developments in spatial panel data models", *Regional Science and Urban Economics*, 40, pp. 255-271.
- LeSage, J.P., and Pace, R.K., 2009, *Introduction to Spatial Econometrics*, Taylor & Francis CRC Press: Boca Raton.

- Madson, A. W., 1916, *The State As Manufacturer and Trader*, T.F. Unwin: London.
- Manera, G., 1963, *La domanda di tabacco in Italia*, Giuffrè: Milan.
- Ministero delle finanze, 1893, *Azienda dei tabacchi. Relazione e bilancio industriale per l'esercizio dal 1 luglio 1892 al 30 giugno 1893*, Rome.
- Mur, J., and Angulo, A., 2006, "The spatial Durbin model and the common factor tests", *Spatial Economic Analysis*, 1, pp. 207-226.
- Mur, J., and Angulo, A., 2009, "Model selection strategies in a spatial setting: some additional results", *Regional Science and Urban Economics* 39, pp. 200-213.
- Parent, O., and LeSage, J.P., 2010, "A spatial dynamic panel model with random effects applied to commuting times", *Transportation Research Part B*, 44, pp. 633-645.
- Pasetti, T., 1906, *Il Monopolio del tabacco in Italia. Cenni storico-statistici*, Ministero delle Finanze. Direzione generale delle privative: Portici.
- Pesaran, M.H., 2004, *General diagnostic tests for cross section dependence in panels*, CESifo Working Paper Series No. 1229. Available at SSRN: <http://ssrn.com/abstract=572504>.
- Pesaran, M.H., 2006, "Estimation and inference in large heterogeneous panels with a multifactor error structure", *Econometrica*, 74, pp. 967-1012.
- Pesaran, M.H., 2015a, "Testing weak cross-sectional dependence in large panels", *Econometric Reviews*, 34, 6-10, pp. 1088-1116.
- Pesaran, M.H., 2015b, *Time Series and Panel Data Econometrics*, Oxford University Press: Oxford.
- Pesaran, M.H., Smith, L.V., and Yamagata, T., 2013, Panel unit root tests in the presence of a multifactor error structure, *Journal of Econometrics*, 175, pp. 94-115.
- Ragioneria generale dello Stato, 1914, *Il bilancio del Regno d'Italia negli esercizi finanziari dal 1862 al 1912-13*, Tipografia dell'Unione Editrice: Rome.
- Rogozinsky, J., 1990, *Smokeless Tobacco in the Western World: 1550-1950*, Praeger: New York, NY.
- Scalzi, F., 1868, *Il consumo di tabacco da fumo in Roma, in attinenza alla salute pubblica*, Stabilimento Tipografico di G. Via: Rome.
- Upton, G.J.G., and Fingleton, B., 1985, *Spatial Data Analysis by Example*, volume 1, Wiley: Chichester.

- Verhelst, B., and Poel, D. van den, 2013, “Deep habits in consumption: a spatial panel analysis using scanner data”, *Empirical Economics*, 47, pp. 959-976.
- Yu, J., Jong, R. de, and Lee, L.-f., 2008, “Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large”, *Journal of Econometrics*, 146, pp. 118-134.
- Yu, J., Jong, R. de, and Lee, L.-f., 2012, “Estimation for spatial dynamic panel data with fixed effects: The case of spatial cointegration”, *Journal of Econometrics*, 167, pp. 16-37.